AD NO. ___________________________
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC-10519

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

BLIND GRID SCORING RECORD NO. 935

SITE LOCATION:
U.S. ARMY YUMA PROVING GROUND

DEMONSTRATOR:
SKY RESEARCH INC.
44 DEAD INDIAN MEMORIAL ROAD
ASHLAND, OREGON 97520

TECHNOLOGY TYPE/PLATFORM:
EMI/MAN PORTABLE VECTOR SENSOR

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

MARCH 2011

Prepared for:
SERDP/ESTCP
MUNITIONS MANAGEMENT
ARLINGTON, VA 22203

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISTRIBUTION UNLIMITED, MARCH 2011.

DISPOSITION INSTRUCTIONS

Destroy this document when no longer needed. Do not return to the originator.

The use of trade names in this document does not constitute an official endorsement or approval of the use of such commercial hardware or software. This document may not be cited for purposes of advertisement.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
March 2011	Final	12-15, 18-22 October 2010

4. TITLE AND SUBTITLE
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE
BLIND GRID SCORING RECORD NO. 935

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
McClung, J. Stephen

5d. PROJECT NUMBER
8-CO-160-UXO-021

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Aberdeen Test Center
ATTN: TEDT-AT-SLE
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER
ATC-10519

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Environmental Command
ATTN: IMAE-RTA
Aberdeen Proving Ground, MD 21010-5401

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Same as item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited, March 2011.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This scoring record documents the efforts of Sky Research Inc. to detect and discriminate inert unexploded ordnance (UXO) utilizing the YPG Standardized UXO Technology Demonstration Site blind grid. This scoring record was coordinated by J. Stephen McClung and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Command, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS
Sky Research Inc., blind grid, man portable EMI Vector System

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
Unclassified	Unclassified	Unclassified			

ACKNOWLEDGMENTS

Authors:

William Burch
Leonard Lombardo
Homeland Defense and Sustainment Division (HDSD)
U.S. Army Aberdeen Test Center
Aberdeen Proving Ground

Rick Fling
Aberdeen Test Support Services (ATSS)
Sverdrup Technology, Inc.
Aberdeen Proving Ground (APG)

Contributors:

J. Stephen McClung
Homeland Defense and Sustainment Division (HDSD)
U.S. Army Aberdeen Test Center
Aberdeen Proving Ground

Christina McClung
Survivability and Lethality (SL) Directorate
U.S. Army Aberdeen Test Center (ATC)
Aberdeen Proving Ground

TABLE OF CONTENTS

PAGE

PAGE

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of unexploded ordnance (UXO) require testing so that their performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland, and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in ordnance and clutter. Testing at these sites is independently administered and analyzed by the Government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments.

The Standardized UXO Technology Demonstration Site Program is a multiagency program spearheaded by the U.S. Army Environmental Command (USAEC). The U.S. Army Aberdeen Test Center (ATC) and the U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) provide programmatic support. The program is being funded and supported by the Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP), and the Army Environmental Quality Technology Program (EQT).

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios that vary targets, geology, clutter, topography, and vegetation.

b. To determine cost, time, and manpower requirements to operate the technology.

c. To determine the demonstrator's ability to analyze survey data in a timely manner and provide prioritized "Target Lists" with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth, geo-referenced data for post-demonstration analysis.

1.2.1 Scoring Methodology

a. The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating

characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}), and those that do not correspond to any known item, termed background alarms.

b. The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the blind grid RESPONSE STAGE, the demonstrator provides the scoring committee with a target response from each and every grid square along with a noise level below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, will include signals both above and below the system noise level.

c. The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such and to reject clutter. For the blind grid DISCRIMINATION STAGE, the demonstrator provides the scoring committee with the output of the algorithms applied in the discrimination-stage processing for each grid square. The values in this list are prioritized based on the demonstrator's determination that a grid square is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking is based on human (subjective) judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance (i.e., that is expected to retain all detected ordnance and rejects the maximum amount of clutter).

d. The demonstrator is also scored on EFFICIENCY and REJECTION RATIO, which measures the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. EFFICIENCY measures the fraction of detected ordnance retained after discrimination, while the REJECTION RATIO measures the fraction of false alarms rejected. Both measures are defined relative to performance at the demonstrator-supplied level below which all responses are considered noise, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

e. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 3.1.1.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response Stage ROC curves:

(1) Probability of Detection (P_d^{res}).

(2) Probability of False Positive (P_{fp}^{res}).

(3) Background Alarm Rate (BAR^{res}) or Probability of Background Alarm (P_{BA}^{res}).

b. Discrimination Stage ROC curves:

(1) Probability of Detection (P_d^{disc}).

(2) Probability of False Positive (P_{fp}^{disc}).

(3) Background Alarm Rate (BAR^{disc}) or Probability of Background Alarm (P_{BA}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False Positive Rejection Rate (R_{fp}).

(3) Background Alarm Rejection Rate (R_{BA}).

d. Other:

(1) Probability of Detection by Size and Depth.

(2) Classification by type (i.e., 20-, 40-, 105-mm, etc.).

(3) Location accuracy.

(4) Equipment setup, calibration time, and corresponding man-hour requirements.

(5) Survey time and corresponding man-hour requirements.

(6) Reacquisition/resurvey time and man-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

1.3 STANDARD AND NONSTANDARD INERT ORDNANCE TARGETS

The standard and nonstandard ordnance items emplaced in the test areas are listed in Table 1. Standardized targets are members of a set of specific ordnance items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are inert ordnance items having properties that differ from those in the set of standardized targets.

TABLE 1. INERT ORDNANCE TARGETS

Standard Type	Nonstandard (NS)
20-mm Projectile M55	20-mm Projectile M55
	20-mm Projectile M97
40-mm Projectile MKII Bodies	40-mm Projectile M813
BDU-28 Submunition	
BLU-26 Submunition	
M42 Submunition	
57-mm Projectile APC M86	
60-mm Mortar M49A3	60-mm Mortar (JPG)
	60-mm Mortar M49
2.75-inch Rocket M230	2.75-inch Rocket M230
	2.75-inch Rocket XM229
81-mm Mortar M374	81-mm Mortar (JPG)
	81-mm Mortar M374
105-mm HEAT Rounds M456	
105-mm Projectile M60	105-mm Projectile M60
155-mm Projectile M483A1	155-mm Projectile M483A
	500-lb Bomb
	M75 Submunition

HEAT = High-explosive antitank.
JPG = Jefferson Proving Ground.

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

POC: Dr. Nicolas Lhomme
541-552-5180

Address: Sky Research Inc.
44 Dead Indian Memorial Road
Ashland, Oregon 97520

2.1.2 System Description (provided by demonstrator)

The MPV is a man-portable, wide-band, time-domain, Electromagnetic Induction (EMI) sensor composed of an array of five receivers that measure all three components of the EM field. The MPV is programmable, such that durations of excitation and time decay recording are adjustable according to survey needs. Development and characterization of the MPV prototype were conducted under SERDP MM-1443 at ERDC-CRREL in Dartmouth, NH. Extensive laboratory tests and preliminary field trials confirmed its potential to extend advanced unexploded ordnance (UXO) classification capabilities to man-portable systems. Field survey with the MPV will be performed with a portable local positioning system. The operating principle consists of locating the origin of the primary field generated by the MPV transmitter coil, acting as a beacon, with a pair of EMI receivers placed at a nearby base station. Position and orientation estimates returned by this beacon system in preliminary field trials were accurate to within 1 centimeter and 1 degree, respectively, out to distances of 5 meters. This system should facilitate advanced discrimination in environments were traditional line-of-sight-based methods fail, e.g., at densely forested sites. A GPS will also be used to locate the base station in global coordinates for comparison with the ground truth.

The MPV-beacon system is to be deployed for UXO detection and the collection of advanced classification-quality data. The MPV features distinct detection and discrimination operating modes with a seamless switch between the two. The former mode consists of dynamic data collection for digital geophysical mapping (DGM). It is based on fast EMI transmit-receive cycles so that the sensor can continuously move (1 millisecond [ms] time decay, similar to Geonics EM-61). The latter is tailored for optimizing data quality and the ensuing target characterization. The sensor is static so that signals can be stacked (averaged to reduce noise); longer EMI cycles are applied to capture variations in time decay rates (e.g., 25 ms, similar to Geonics EM-63). The MPV interactive user interface has real-time monitoring and feedback capabilities on data quality, spatial coverage and other key features (signal intensity, time decay, secondary targets, and presence of magnetic soil). These features assist the field operator in efficient data collection, so that detection and discrimination data can be collected as part of the same survey, thus limiting the need to revisit an anomaly for further characterization.

Figure 1. MPV/EMI.

2.1.3 Data Processing Description (provided by demonstrator)

Detection data and static, cued interrogation data will be collected as part of the same survey. Detection data will be interpreted in real-time in the field through the sensor operating interface to assess whether an anomaly requires immediate cued interrogation, and recorded to generate a post-survey map of the studied area. Cued interrogation data will be inverted at a later stage to classify anomalies and identify potential UXO.

Target selection:

- The MPV records both detection data and cued interrogation data in .tem binary files that are converted to .csv files (ASCII) for post-processing. The raw data are stacked and gated before recording without any additional pre-processing;
- Post-survey data processing is performed in MatLab; the final data are stored in .mat files binary format;
- Targets are detected when the amplitude of the vertical-component data exceeds a given threshold and a change in sign is observed in the horizontal-component data. This heuristic target detection algorithm takes advantage of having three-orthogonal-component receivers, given that EMI theory predicts peak amplitude of EM field vertical component above a compact metallic body (possibility of two peaks for near-horizontal target), and change in sign in both horizontal components of EM field right above that target. Target selection will be manually validated by the expert geophysicist on site (automatic target detection process is to be developed at later stage and trained with field data).

- Detection threshold for the vertical-component data will be preset at the start of the deployment, then re-assessed and refined on site to reflect experience and understanding of environmental field conditions. The initial threshold will be inferred from shakedown tests and simulations prior to visiting YPG. Detection rate at the YPG CALIBRATION LANES will provide training data to update the threshold. The actual choice of threshold amplitude is non-essential to detection performance, as horizontal-component data bring additional constraints.

<u>Parameter estimation:</u>

Cued interrogation data collected in static mode (long time decay) will be inverted to recover target location, orientation, and the principle components of the magnetic polarizability tensor.

- Model parameters will be obtained by solving the least-square fit to dipole model using the Sky Research-University of British Columbia software UXOLab. Magnetic charge models (Barrowes and Schubitidze) will also be tested at Dartmouth to assess whether there are instances where superior performance warranties the computational overhead of using such models;
- Discrimination with dipole model inversions will be performed by extracting parameters from principle components of magnetic polarizability tensor as a function of time, as these relate to size, aspect ratio and material properties of compact metallic body. Empirical evidence gained through previous studies for SERDP and ESTCP suggest that amplitude and time decay rate of these components are powerful discrimination factors, as clutter items generally present faster time decay rate or smaller size than UXO. When data quality supports reliable recovery of the tertiary polarizability component a shape factor will be utilized to separate oblong from plate-like targets;
- Target depth and geographic location will be obtained by combining predicted location with GPS reading at beacon base station (OPEN FIELD survey);
- Fit quality, estimated through the difference between observed and predicted data, and their correlation coefficient, will be utilized to separate anomalies that can reliably be discriminated from those that cannot be analyzed and therefore should be excavated;

Environmental noise and positional errors can degrade inversion results. The former contribution is estimated by identifying areas with pure background signal to provide a base level noise; the latter depends on distance from base station but its effect is dampened by perfect positioning between all five receiver cubes at each survey locations. Positional error is accounted through a tunable data percent error. Sensitivity of inversion results to data error estimates will be assessed by measuring the variability in recovered parameters at CALIBRATION LANES, where UXO types are known. This process also helps determine the degree of reliability of inversion results for various noise levels and target size. Besides, if environmental noise originates from highly magnetic soil, models developed under SERDP MM-1573 (Pasion) will be employed to compensate for soil effect on EMI sensor.

Classification:

We propose to test two different strategies for statistical classification:

- Feature-based classification utilizing size and decay rate properties of the magnetic polarizabilities, in addition to shape based measures (i.e., properties of secondary polarizabilities). Choice of a statistical classifier (Probabilistic Neural Network, Support Vector Machine…) will depend on distribution of features (clusters) in model space. The associated munition/non-munition threshold will be determined by assessing sensitivity of UXO probability to training data;
- A library based method will be used to rank anomalies. Anomalies that can be modeled by targets in the library will be ranked higher. Probability of UXO will be defined as the smallest difference between polarizability components of a target relative to library items;

A ranked-priority dig list will be provided for each method. Anomalies with poor reliability measures (poor data fit, high noise) will be recommended for excavation.

Training:

Analysis of the CALIBRATION LANES data will help:

- Determine detection threshold;
- Define data error percent estimate by measuring its effect on recovered model parameters relative to ground truth;
- Determine the sensitivity of classification features (cluster spread);
- Define the optimum classifier and threshold by measuring their sensitivity to random subsets of the training data;
- Build a library of polarizability components for known ordnance.

2.1.4 Data Submission Format

Data were submitted for scoring in accordance with data submission protocols outlined in the Standardized UXO Technology Demonstration Site Handbook. These submitted data are not included in this report in order to protect ground truth information.

2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC) (provided by demonstrator)

Overview of Quality Control (QC):

The MPV sensor interface provides numerous QC tools to detect potential loss of positional accuracy and monitor system performance by displaying the recorded signal almost in real-time, allowing detection of receiver malfunction and environmental noise intensity, and monitoring battery life. Sensor drift should not occur owing to stable EM components design. Local positional accuracy will also be controlled by periodically placing the MPV at reference locations with known separation.

Detection will also be part of a QC procedure. The recorded detection data will be utilized to generate a post-survey map of the studied area. This process is particularly important because the data will be collected as individual cells. Including all cells as part of a unique map facilitates verification of complete spatial coverage (no gaps). This will also serve as a unified control on target picking to ensure that no potential target could have been missed, for instance by re-assessing detection threshold.

Overview of Quality Assurance (QA):

The survey design will ensure 100% coverage of every studied area by establishing partially overlapping survey lines (0.40 m line spacing with 0.50 m sensor head diameter) and high sampling rate (0.1 s) along lines. Navigation accuracy will be ensured by staying within 5 m of the beacon base station, with maximum cell size of 4 m x 4 m (maximum in calibration lanes).

2.1.6 Additional Records

The following record(s) by this vendor can be accessed via the Internet as Microsoft Word documents at www.uxotestsites.org.

2.2 YPG SITE INFORMATION

2.2.1 Location

YPG is located adjacent to the Colorado River in the Sonoran Desert. The UXO Standardized Test Site is located south of Pole Line Road and east of the Countermine Testing and Training Range. The open field range, calibration grid, blind grid, mogul area, and desert extreme area comprise the 350- by 500-meter general test site area. The open field site is the largest of the test sites and measures approximately 200 by 350 meters. To the east of the open field range are the calibration and blind test grids that measure 30 by 40 meters and 40 by 40 meters, respectively. South of the open field is the 135- by 80-meter mogul area consisting of a sequence of man-made depressions. The desert extreme area is located southeast of the open field site and has dimensions of 50 by 100 meters. The desert extreme area, covered with desert-type vegetation, is used to test the performance of different sensor platforms in a more severe desert conditions/environment.

2.2.2 Soil Type

Soil samples were collected at the YPG UXO Standardized Test Site by ERDC to characterize the shallow subsurface (< 3 m). Both surface grab samples and continuous soil borings were acquired. The soils were subjected to several laboratory analyses, including sieve/hydrometer, water content, magnetic susceptibility, dielectric permittivity, X-ray diffraction, and visual description.

Two soil complexes are present within the site: Riverbend-Carrizo and Cristobal-Gunsight. The Riverbend-Carrizo complex is composed of mixed stream alluvium, whereas the Cristobal-Gunsight complex is derived from fan alluvium. The Cristobal-Gunsight complex covers the majority of the site. Most of the soil samples were classified as either a sandy loam or loamy sand, with most samples containing gravel-size particles. All samples had a measured water content less than 7 percent, except for two that contained 11-percent moisture. The majority of soil samples had water content between 1 and 2 percent. Samples containing more than 3 percent were generally deeper than 1 meter.

An X-ray diffraction analysis on four soil samples indicated a basic mineralogy of quartz, calcite, mica, feldspar, magnetite, and some clay. The presence of magnetite imparted a moderate magnetic susceptibility, with volume susceptibilities generally greater than 100 by 105 SI.

For more details concerning the soil properties at the YPG test site, go to www.uxotestsites.org on the Web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at YPG is included in Table 2.

TABLE 2. TEST SITE AREAS

Area	Description
Calibration grid	Contains the 15 standard ordnance items buried in six positions at various angles and depths to allow demonstrator equipment calibration.
Blind grid	Contains 400 grid cells in a 0.16-hectare (0.39-acre) site. The center of each grid cell contains ordnance, clutter, or nothing.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (12-15, 18-22 October 2010)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total number of hours operated at each site are summarized in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	Number of Hours
Calibration lanes	18.61
Blind grid	34.20

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

A YPG weather station located approximately 1 mile west of the test site was used to record average temperature and precipitation on a half-hour basis for each day of operation. The temperatures listed in Table 4 represent the average temperature during field operations from 0700 to 1700 hours, while precipitation data represent a daily total amount of rainfall. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 2010	Average Temperature, °F	Total Daily Precipitation, in.
October 12	87.1	0.00
October 13	91.6	0.00
October 14	94.9	0.00
October 15	91.4	0.00
October 18	80.4	0.00
October 19	77.2	0.00
October 20	69.5	0.00
October 21	71.3	0.00
October 22	75.9	0.00

3.3.2 Field Conditions

Sky Research surveyed the blind grid 14-15, 18-20, 22 October 2010. The weather was warm, and the field was dry during the survey.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: calibration, mogul, open field, and desert extreme areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are included in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization, daily equipment preparation and breakdown. A four-person crew took 3 hours and 24 minutes to perform the initial setup and mobilization. There was 2 hours eight minutes of daily equipment preparation and 1 hour 38 minutes of equipment breakdown.

3.4.2 Calibration

Sky Research spent a total of 18 hours and 37 minutes in the calibration lanes, of which 9 hours and 27 minutes were spent collecting data.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, demonstration site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor costs (section 5) except for downtime due to demonstration site issues. Demonstration site issues, while noted in the daily log, are considered non-chargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total site survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment data checks and maintenance activities accounted for 36 minutes of site usage time. These activities included changing out batteries and performing routine data checks to ensure the data were being properly recorded/collected. Sky spent an additional 3 hours 22 minutes for breaks and lunches.

3.4.3.2 Equipment failure or repair. 5 hours 27 minutes was needed to resolve equipment failures that occurred while surveying the blind grid. A positional sensing wire broke and to be replaced. Two other small delays occurred with this wire.

3.4.3.3 Weather. A 60 minute weather delay occurred on 20 October 2010. The delay was due to rain.

3.4.4 Data Collection

Sky spent a total time of 34 hours and 12 minutes in the blind grid area, of which 19 hours 40 minutes was spent collecting data.

3.4.5 Demobilization

The Sky survey crew went on to conduct a demonstration of the site. Therefore, demobilization did not occur until 22 October 2010. On that day, it took the crew 50 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

Sky submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data were provided 23 November 2010.

3.6 DEMONSTRATOR'S FIELD PERSONNEL

Nicolas Lhomme
Jon Jacobsen
Benjamin Barrowes
David George

3.7 DEMONSTRATOR'S FIELD SURVEYING METHOD

Sky Research surveyed the blind grid in a linear fashion, in a north-to-south and east-to-west direction.

3.8 SUMMARY OF DAILY LOGS

Daily logs captured all field activities during this demonstration and are located in Appendix D. Activities pertinent to this specific demonstration are indicated in highlighted text.

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL ORDNANCE CATEGORIES

The probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive are shown in Figure 2. Both probabilities plotted against their respective background alarm rate are shown in Figure 3. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

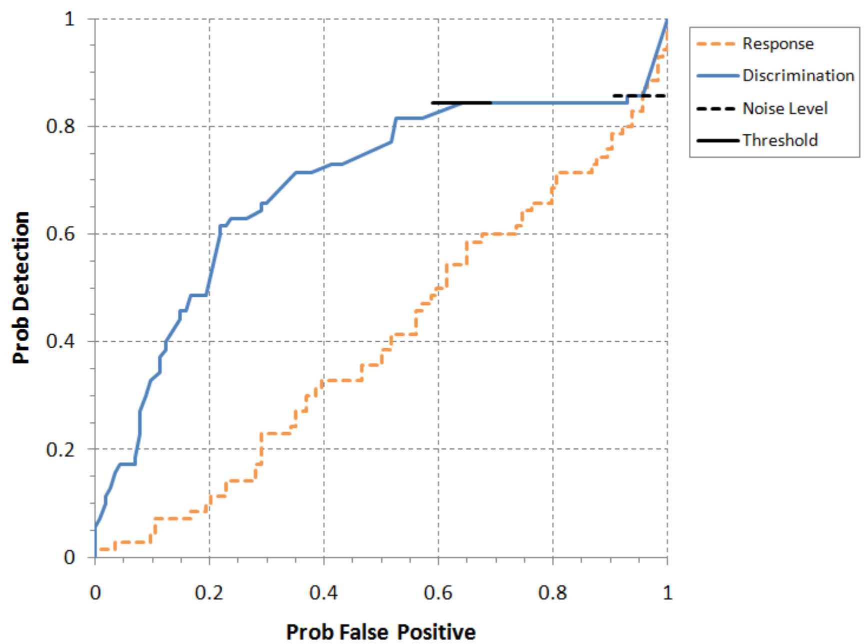

Figure 2. MPV/handheld blind grid probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

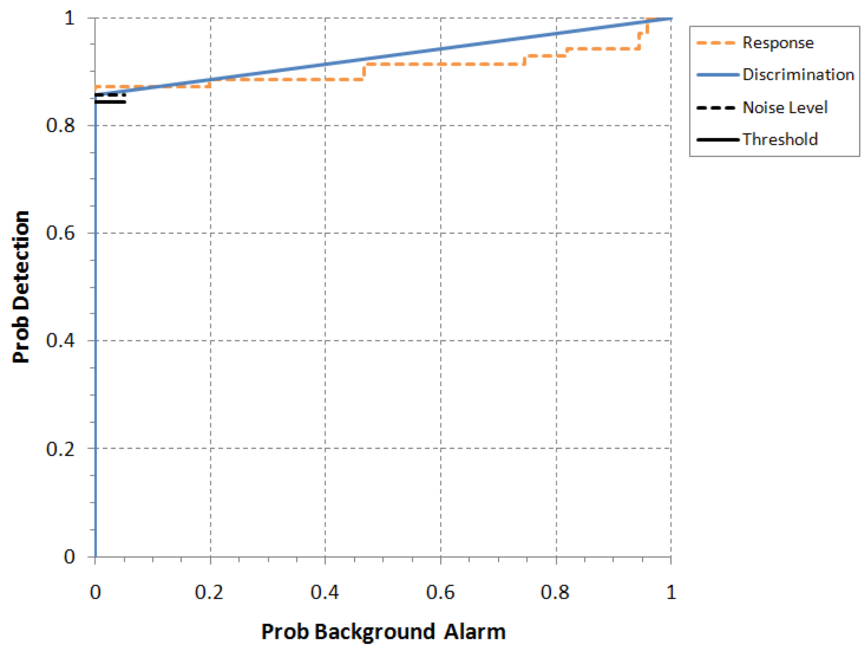

Figure 3. MPV /handheld blind grid probability of detection for response and discrimination stages versus their respective probability of background alarm over all ordnance categories combined.

4.2 ROC CURVES USING ORDNANCE LARGER THAN 20 MM

The probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive when only targets larger than 20 mm are scored are shown in Figure 4. Both probabilities plotted against their respective background alarm rate are shown in Figure 5. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

NA

Figure 4. MPV /handheld blind grid probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

NA

Figure 5. MPV /handheld blind grid probability of detection for response and discrimination stages versus their respective probabilities of background alarm for all ordnance larger than 20 mm.

4.3 PERFORMANCE SUMMARIES

Results for the open field test broken out by size, depth, and nonstandard ordnance are presented in Table 5 (for cost results, see section 5). Results by size and depth include both standard and nonstandard ordnance. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range (see app A for size definitions). The results are relative to the number of ordnance items emplaced. Depth is measured from the geometric center of anomalies.

The RESPONSE STAGE results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the DISCRIMINATION STAGE are derived from the demonstrator's recommended threshold for optimizing UXO field cleanup by minimizing false digs and maximizing ordnance recovery. The lower 90-percent confidence limit on probability of detection and probability of false positive was calculated assuming that the number of detections and false positives are binomially distributed random variables. All results in Table 5 have been rounded to protect the ground truth. However, lower confidence limits were calculated using actual results.

TABLE 5. SUMMARY OF BLIND GRID RESULTS FOR THE MPV/HANDHELD

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.85	0.90	0.80	0.95	0.75	0.80	1.00	0.90	0.00
P_d Low 90% Conf	0.79	0.79	0.68	0.89	0.59	0.58	0.94	0.76	0.00
P_d Upper 90% Conf	0.91	0.94	0.91	1.00	0.86	0.92	1.00	0.96	0.28
P_{fp}	0.95	-	-	-	-	-	0.95	0.95	NA
P_{fp} Low 90% Conf	0.92	-	-	-	-	-	0.92	0.83	NA
P_{fp} Upper 90% Conf	0.98	-	-	-	-	-	0.99	0.98	NA
P_{ba}	0.00	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.85	0.85	0.80	0.95	0.75	0.80	0.95	0.90	0.00
P_d Low 90% Conf	0.77	0.77	0.68	0.85	0.59	0.58	0.90	0.76	0.00
P_d Upper 90% Conf	0.90	0.93	0.91	0.98	0.86	0.92	1.00	0.96	0.28
P_{fp}	0.65	-	-	-	-	-	0.60	0.75	0.00
P_{fp} Low 90% Conf	0.58	-	-	-	-	-	0.52	0.64	NA
P_{fp} Upper 90% Conf	0.70	-	-	-	-	-	0.67	0.87	NA
P_{ba}	0.00	-	-	-	-	-	-	-	-

Response Stage Noise Level: 14.
Recommended Discrimination Stage Threshold: 24.

Note: The recommended discrimination stage threshold values are provided by the demonstrator.

4.4 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table 6.

TABLE 6. EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	0.98	0.33	NA
With No Loss of Pd	1.00	0.03	NA

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (table 7). Correct type examples include 20-mm projectile, 105-mm HEAT Projectile, and 2.75-inch Rocket. A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20 mm, 105 H, and 2.75 in., respectively.

TABLE 7. CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS UXO

Size	Percentage Correct
Small	0.88
Medium	0.57
Large	0.57
Overall	0.71

4.5 LOCATION ACCURACY

The mean location error and standard deviations appear in Table 8. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the blind grid, only depth errors are calculated because (X, Y) positions are known to be the centers of each grid square.

TABLE 8. MEAN LOCATION ERROR AND STANDARD DEVIATION (M)

	Mean	Standard Deviation
Depth	-0.01	0.09

SECTION 5. ON-SITE LABOR COSTS

A standardized estimate for labor costs associated with this effort was calculated as follows: the first person at the test site was designated supervisor, the second person was designated data analyst, and the third and following personnel were considered field support. Standardized hourly labor rates were charged by title: supervisor at $95.00/hour, data analyst at $57.00/hour, and field support at $28.50/hour.

Government representatives monitored on-site activity. All on-site activities were grouped into one of ten categories: initial setup/mobilization, daily setup/stop, calibration, data collection, downtime due to break/lunch, downtime due to equipment failure, downtime due to equipment/data checks or maintenance, downtime due to weather, downtime due to demonstration site issue, or demobilization. See Appendix D for the daily activity log. See section 3.4 for a summary of field activities.

The standardized cost estimate associated with the labor needed to perform the field activities is presented in Table 9. Note that calibration time includes time spent in the calibration lanes as well as field calibrations. Site survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to equipment/data checks or maintenance, downtime due to failure, and downtime due to weather.

TABLE 9. ON-SITE LABOR COSTS

	No. People	Hourly Wage	Hours	Cost
Initial setup				
Supervisor	1	$95.00	3.40	$323.00
Data analyst	1	57.00	3.40	193.80
Field support	2	28.50	3.40	193.80
Subtotal				**$710.60**
Calibration				
Supervisor	1	$95.00	18.61	$1767.95
Data analyst	1	57.00	18.61	1060.77
Field support	2	28.50	18.61	1060.77
Subtotal				**$3889.49**
Site survey				
Supervisor	1	$95.00	34.20	$3249.00
Data analyst	1	57.00	34.20	1949.40
Field support	0	28.50		
Subtotal				**$5198.40**

See notes at end of table.

TABLE 9 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
Demobilization				
Supervisor	1	$95.00	0.83	$78.85
Data analyst	1	57.00	0.83	47.31
Field support		28.50		
Subtotal				**$126.16**
Total				**$9924.65**

Notes: Calibration time includes time spent in the calibration lanes as well as calibration before each data run.
Site survey time includes daily setup/stop time, collecting data, breaks/lunch, and downtime due to system maintenance, failure, and weather.

SECTION 6.   COMPARISON OF RESULTS TO DATE

No comparisons to date.

SECTION 7. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. If multiple declarations lie within R_{halo} of any item (clutter or ordnance), the declaration with the highest signal output within the R_{halo} will be utilized. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the length of the ordnance plus 1 meter.

Small Ordnance: Caliber of ordnance less than or equal to 40 mm (includes 20-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Ordnance: Caliber of ordnance greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75 in. Rocket, 81-mm mortar).

Large Ordnance: Caliber of ordnance greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, 155-mm projectile, 500-pound bomb).

Shallow: Items buried less than 0.3 meter below ground surface.

Medium: Items buried greater than or equal to 0.3 meter and less than 1 meter below ground surface.

Deep: Items buried greater than or equal to 1 meter below ground surface.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the blind grid test area.

Discrimination Stage Threshold: The demonstrator selected threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}) and those that do not correspond to any known item, termed background alarms.

The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the RESPONSE STAGE, the demonstrator provides the scoring committee with the location and signal strength of all anomalies that the demonstrator has deemed sufficient to warrant further investigation and/or processing as potential emplaced ordnance items. This list is generated with minimal processing (e.g., this list will include all signals above the system noise threshold). As such, it represents the most inclusive list of anomalies.

The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such, and to reject clutter. For the same locations as in the RESPONSE STAGE anomaly list, the DISCRIMINATION STAGE list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide "optimum" system performance, (i.e., that retains all the detected ordnance and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open Field only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to non-ordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

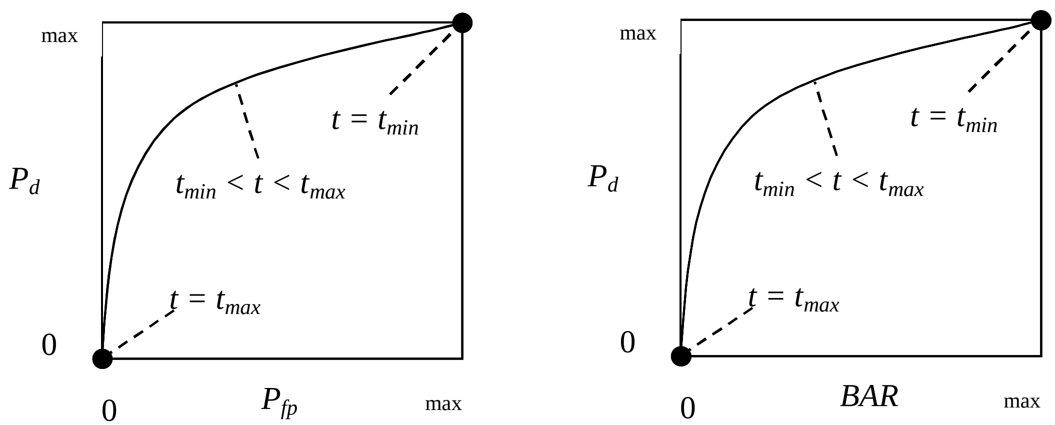

Figure A-1. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a pre-determined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the blind grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$; Measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage tmin) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$; Measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage tmin). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})])$.

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 3).

A 2 x 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X’s system is significantly degraded by the more challenging terrain feature introduced. The test statistic of the 2 x 2 contingency table is the

Chi-square distribution with one degree of freedom. Since an association between the more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A significance level of 0.05 is chosen which sets a critical decision limit of 2.71 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, the proportions are considered to be significantly different.

Standardized UXO Technology Demonstration Site examples, where blind grid results are compared to those from the open field and open field results are compared to those from one of the scenarios, follow. It should be noted that a significant result does not prove a cause and effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three progressively more difficult areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind grid	Open field	Moguls
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 ordnance out of 10 emplaced were correctly discriminated as such in open field-testing. Those four values are used to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 2.71, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{disc}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the smaller discrimination stage detection rate is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the ability of demonstrator X to correctly discriminate seems to have been degraded by the mogul terrain relative to results from the flat open field using the same system.

APPENDIX B. DAILY WEATHER LOGS

Date, 2010	Time, EST	Average Temperature, ^{o}F	Average Precipitation, in.
12 October	0700	67.9	0.00
	0800	70.5	0.00
	0900	77.5	0.00
	1000	84.5	0.00
	1100	86.0	0.00
	1200	89.4	0.00
	1300	93.5	0.00
	1400	95.5	0.00
	1500	97.2	0.00
	1600	98.0	0.00
	1700	98.3	0.00
13 October	0700	70.3	0.00
	0800	73.6	0.00
	0900	81.3	0.00
	1000	87.8	0.00
	1100	92.3	0.00
	1200	96.8	0.00
	1300	99.6	0.00
	1400	100.5	0.00
	1500	101.7	0.00
	1600	102.3	0.00
	1700	102.0	0.00
14 October	0700	82.5	0.00
	0800	82.9	0.00
	0900	87.7	0.00
	1000	88.9	0.00
	1100	91.7	0.00
	1200	94.1	0.00
	1300	95.3	0.00
	1400	96.0	0.00
	1500	98.9	0.00
	1600	99.6	0.00
	1700	99.7	0.00

Date, 2010	Time, EST	Average Temperature, °F	Average Precipitation, in.
15 October	0700	78.6	0.00
	0800	79.3	0.00
	0900	83.6	0.00
	1000	86.9	0.00
	1100	88.5	0.00
	1200	91.2	0.00
	1300	92.0	0.00
	1400	93.6	0.00
	1500	94.4	0.00
	1600	95.6	0.00
	1700	95.0	0.00
18 October	0700	66.6	0.00
	0800	69.9	0.00
	0900	74.7	0.00
	1000	76.8	0.00
	1100	77.9	0.00
	1200	80.0	0.00
	1300	82.4	0.00
	1400	82.0	0.00
	1500	83.7	0.00
	1600	82.6	0.00
	1700	82.8	0.00
19 October	0700	68.1	0.00
	0800	71.0	0.00
	0900	73.5	0.00
	1000	76.3	0.00
	1100	77.3	0.00
	1200	80.0	0.00
	1300	79.8	0.00
	1400	80.3	0.00
	1500	73.5	0.00
	1600	72.2	0.00
	1700	72.7	0.00
20 October	0700	60.8	0.00
	0800	62.3	0.00
	0900	65.9	0.00
	1000	67.0	0.00
	1100	69.1	0.00
	1200	72.4	0.00
	1300	72.8	0.00
	1400	71.8	0.00
	1500	65.3	0.00
	1600	68.6	0.00
	1700	69.2	0.00

Date, 2010	Time, EST	Average Temperature, °F	Average Precipitation, in.
21 October	0700	54.5	0.00
	0800	59.6	0.00
	0900	65.8	0.00
	1000	66.1	0.00
	1100	68.5	0.00
	1200	71.4	0.00
	1300	72.9	0.00
	1400	74.9	0.00
	1500	76.2	0.00
	1600	76.6	0.00
	1700	76.2	0.00
22 October	0700	61.5	0.00
	0800	64.1	0.00
	0900	70.6	0.00
	1000	72.2	0.00
	1100	74.4	0.00
	1200	76.6	0.00
	1300	76.5	0.00
	1400	77.9	0.00
	1500	79.1	0.00
	1600	79.4	0.00
	1700	78.5	0.00

APPENDIX C. SOIL MOISTURE

Date: 12 October 2010			
Times: NA and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	-	38.0
	6 to 12	-	36.2
	12 to 24	-	8.1
	24 to 36	-	2.5
	36 to 48	-	8.4
Mogul field	0 to 6	-	38.0
	6 to 12	-	11.1
	12 to 24	-	7.1
	24 to 36	-	3.8
	36 to 48	-	3.9
Desert Extreme area	0 to 6	-	11.1
	6 to 12	-	36.2
	12 to 24	-	7.1
	24 to 36	-	3.8
	36 to 48	-	3.9

Date: 13 October 2010			
Times: 0900 and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	38.0	11.1
	6 to 12	36.2	36.2
	12 to 24	54.3	7.1
	24 to 36	11.1	3.8
	36 to 48	11.1	3.9
Mogul field	0 to 6	11.1	38.0
	6 to 12	36.2	36.2
	12 to 24	7.1	67.4
	24 to 36	95.9	3.8
	36 to 48	6.4	3.9
Desert Extreme area	0 to 6	38.0	11.1
	6 to 12	36.2	36.2
	12 to 24	7.1	7.1
	24 to 36	95.9	6.3
	36 to 48	6.4	3.9

Date: 14 October 2010			
Times: 0700 and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	38.0	11.1
	6 to 12	36.2	36.2
	12 to 24	7.1	11.8
	24 to 36	3.8	8.3
	36 to 48	3.9	5.7
Mogul field	0 to 6	33.0	0.3
	6 to 12	36.2	36.2
	12 to 24	7.1	7.1
	24 to 36	3.8	7.4
	36 to 48	3.9	9.8
Desert Extreme area	0 to 6	11.1	38.0
	6 to 12	36.2	36.2
	12 to 24	7.1	7.1
	24 to 36	3.8	3.8
	36 to 48	3.9	3.9

Date: 15 October 2010			
Times: 0700 and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	2.1	9.2
	6 to 12	7.7	36.2
	12 to 24	6.8	9.8
	24 to 36	6.0	3.8
	36 to 48	9.2	3.9
Mogul field	0 to 6	38.0	39.2
	6 to 12	36.2	36.2
	12 to 24	7.1	7.1
	24 to 36	6.6	3.8
	36 to 48	9.8	3.9
Desert Extreme area	0 to 6	11.1	11.1
	6 to 12	36.2	36.2
	12 to 24	0.7	7.1
	24 to 36	5.9	3.9
	36 to 48	3.8	3.8

Date: 18 October 2010			
Times: 1000 and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	3.0	11.1
	6 to 12	5.4	36.2
	12 to 24	7.4	7.1
	24 to 36	3.1	3.8
	36 to 48	4.3	3.9
Mogul field	0 to 6	38.0	38.0
	6 to 12	36.2	36.2
	12 to 24	7.1	67.4
	24 to 36	6.4	3.8
	36 to 48	8.6	3.9
Desert Extreme area	0 to 6	38.0	11.1
	6 to 12	36.2	36.2
	12 to 24	0.1	7.1
	24 to 36	3.8	6.3
	36 to 48	3.9	3.9

Date: 19 October 2010			
Times: 0700 and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	5.2	2.1
	6 to 12	1.5	7.7
	12 to 24	6.2	6.8
	24 to 36	3.3	6.0
	36 to 48	9.0	9.2
Mogul field	0 to 6	0.6	38.0
	6 to 12	36.2	36.2
	12 to 24	7.1	7.1
	24 to 36	8.3	6.6
	36 to 48	8.5	9.8
Desert Extreme area	0 to 6	11.1	11.1
	6 to 12	36.2	36.2
	12 to 24	0.6	0.7
	24 to 36	5.8	5.9
	36 to 48	3.9	3.8

Date: 20 October 2010			
Times: 0900 and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	1.3	11.1
	6 to 12	4.6	36.2
	12 to 24	6.5	7.1
	24 to 36	3.7	3.8
	36 to 48	8.5	3.9
Mogul field	0 to 6	1.6	38.0
	6 to 12	6.0	36.2
	12 to 24	1.0	67.4
	24 to 36	7.6	3.8
	36 to 48	9.2	3.9
Desert Extreme area	0 to 6	5.2	11.1
	6 to 12	10.7	36.2
	12 to 24	1.0	7.1
	24 to 36	5.5	6.3
	36 to 48	3.9	3.9

Date: 21 October 2010			
Times: 0700 and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	4.9	2.7
	6 to 12	3.5	4.9
	12 to 24	5.5	7.2
	24 to 36	3.1	2.5
	36 to 48	8.9	9.0
Mogul field	0 to 6	38.0	4.4
	6 to 12	36.2	2.9
	12 to 24	2.1	1.1
	24 to 36	7.0	3.8
	36 to 48	8.9	3.9
Desert Extreme area	0 to 6	11.1	11.1
	6 to 12	11.0	36.2
	12 to 24	0.8	0.6
	24 to 36	5.0	3.8
	36 to 48	3.9	3.9

Date: 22 October 2010			
Times: 0800 and NA			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration area	0 to 6	38.0	-
	6 to 12	36.2	-
	12 to 24	6.4	-
	24 to 36	2.7	-
	36 to 48	8.4	-
Mogul field	0 to 6	38.0	-
	6 to 12	36.2	-
	12 to 24	3.4	-
	24 to 36	7.1	-
	36 to 48	9.0	-
Desert Extreme area	0 to 6	11.1	-
	6 to 12	36.2	-
	12 to 24	0.7	-
	24 to 36	34.1	-
	36 to 48	3.9	-

APPENDIX D. DAILY ACTIVITY LOGS

Date, 10	No. of People	Area-Tested	Status Start Time	Status Stop Time	Duration, min.	Operational Status	Operational Status - Comments	Track Method	Pattern	Field Conditions	
12-Oct	4	CALIBRATION LANES	750	1114	204	INITIAL SET-UP	Initial Mobilization	NA	Linear	Sunny	Cool
12-Oct	4	CALIBRATION LANES	1114	1202	48	BREAK/LUNCH	lunch	GPS	Linear	Sunny	Cool
12-Oct	4	CALIBRATION LANES	1202	1439	157	COLLECTING DATA	Collecting Data, North - South, West - East	NA	NA	Sunny	Warm
12-Oct	4	CALIBRATION LANES	1439	1500	21	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny	Warm
13-Oct	4	CALIBRATION LANES	630	758	88	DAILY START, STOP	Setup of equipment	NA	NA	Sunny	Cool
13-Oct	4	CALIBRATION LANES	758	930	92	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
13-Oct	4	CALIBRATION LANES	930	957	27	BREAK/LUNCH	break	NA	NA	Sunny	Warm
13-Oct	4	CALIBRATION LANES	957	1440	283	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Warm
13-Oct	4	CALIBRATION LANES	1440	1458	18	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny	Warm
14-Oct	4	CALIBRATION LANES	625	658	33	DAILY START, STOP	Setup of equipment	NA	NA	Sunny	Cool
14-Oct	4	CALIBRATION LANES	658	735	37	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
14-Oct	4	CALIBRATION LANES	735	746	11	DOWNTIME DUE TO EQUIP MAINT/CHECK	Verifying data	NA	NA	Sunny	Cool
14-Oct	4	CALIBRATION LANES	746	1140	234	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Warm
14-Oct	4	CALIBRATION LANES	1140	1215	35	BREAK/LUNCH	lunch	NA	NA	Sunny	Warm
14-Oct	4	BLIND TEST GRID	1215	1420	125	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Warm
14-Oct	4	BLIND TEST GRID	1420	1445	25	DAILY START, STOP	Break down end of day	NA	NA	Sunny	Warm
15-Oct	3	BLIND TEST GRID	631	655	24	DAILY START, STOP	Setup of equipment	NA	NA	Sunny	Cool
15-Oct	3	BLIND TEST GRID	655	725	30	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
15-Oct	3	BLIND TEST GRID	725	801	36	DOWNTIME DUE TO EQUIP MAINT/CHECK	Verifying data	NA	NA	Sunny	Cool
15-Oct	3	BLIND TEST GRID	801	1310	309	DOWNTIME DUE TO EQUIPMENT FAILURE	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
15-Oct	2	BLIND TEST GRID	1310	1335	25	BREAK/LUNCH	lunch	NA	NA	Sunny	Warm
15-Oct	2	BLIND TEST GRID	1335	1440	65	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Warm

Date, 10	No. of People	Area-Tested	Status Start Time	Status Stop Time	Duration, min.	Operational Status	Operational Status - Comments	Track Method	Pattern	Field Conditions	
15-Oct	2	BLIND TEST GRID	1440	1455	15	DAILY START, STOP	Breakdown end of day, At the of day while packing up the system, a Wire connector on the experimental positioning sensor was broken off	GPS	Linear	Sunny	Warm
18-Oct	2	CALIBRATION LANES	1001	1034	33	DAILY START, STOP	Setup of equipment, Late start due to waiting on pick up of new wire for experimental positioning sensor from FedEx	NA	NA	Sunny	Cool
18-Oct	2	BLIND TEST GRID	1034	1148	74	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
18-Oct	2	BLIND TEST GRID	1148	1159	11	DOWNTIME DUE TO EQUIPMENT FAILURE	Reported 2nd wire from experimental positioning sensor is broken also, will attempt to have fixed after end of day	NA	NA	Sunny	Cool
18-Oct	2	BLIND TEST GRID	1159	1210	11	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
18-Oct	2	BLIND TEST GRID	1210	1225	23580	BREAK/LUNCH	lunch	NA	NA	Sunny	Warm
18-Oct	2	BLIND TEST GRID	1225	1425	120	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Warm
18-Oct	2	BLIND TEST GRID	1425	1432	7	DOWNTIME DUE TO EQUIPMENT FAILURE	Reported Orientation sensor on hand held unit not responding, will trouble shoot at end of day	NA	NA	Sunny	Warm
18-Oct	2	BLIND TEST GRID	1432	1450	18	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny	Warm
19-Oct	2	BLIND TEST GRID	628	700	32	DAILY START, STOP	Setup of equipment	NA	NA	Sunny	Cool
19-Oct	2	BLIND TEST GRID	700	925	145	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
19-Oct	2	BLIND TEST GRID	925	938	13	BREAK/LUNCH	Break, water	NA	NA	Sunny	Cool
19-Oct	2	BLIND TEST GRID	938	1210	152	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool

Date, 10	No. of People	Area-Tested	Status Start Time	Status Stop Time	Duration, min.	Operational Status	Operational Status - Comments	Track Method	Pattern	Field Conditions	
19-Oct	2	BLIND TEST GRID	1210	1435	145	BREAK/LUNCH	Lunch	NA	NA	Sunny	Cool
19-Oct	2	BLIND TEST GRID	1435	1455	20	DAILY START, STOP	Breakdown End of Day	NA	NA	Sunny	Cool
20-Oct	2	BLIND TEST GRID	625	738	73	DAILY START, STOP	Setup of equipment	NA	NA	Sunny	Cool
20-Oct	2	BLIND TEST GRID	738	1100	202	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
20-Oct	2	BLIND TEST GRID	1100	1124	24	BREAK/LUNCH	Lunch	NA	NA	Sunny	cool
20-Oct	2	BLIND TEST GRID	1124	1245	81	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
20-Oct	2	BLIND TEST GRID	1245	1345	60	WEATHER ISSUE	Rain, Downloading Data	NA	NA	Sunny	Cool
20-Oct	2	BLIND TEST GRID	1345	1500	75	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
20-Oct	2	BLIND TEST GRID	1500	1520	20	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny	Cool
21-Oct	2	DESERT EXTREME	625	738	73	DAILY START, STOP	Setup of equipment	NA	NA	Sunny	Cool
21-Oct	2	DESERT EXTREME	738	1131	233	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
21-Oct	2	DESERT EXTREME	1131	1220	49	BREAK/LUNCH	lunch	GPS	Linear	Sunny	Cool
21-Oct	2	DESERT EXTREME	1220	1400	100	DOWNTIME DUE TO EQUIPMENT FAILURE	Unable to sync GPS with system	GPS	Linear	Sunny	Cool
21-Oct	2	DESERT EXTREME	1400	1440	40	DAILY START, STOP	Breakdown End of Day	NA	NA	Sunny	Warm
22-Oct	2	DESERT EXTREME	635	800	85	DAILY START, STOP	Setup of Equipment	NA	NA	Sunny	Cool
22-Oct	2	BLIND TEST GRID	800	940	100	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
22-Oct	2	DESERT EXTREME	940	1010	30	DOWNTIME DUE TO EQUIP MAINT/CHECK	setting up equipment for Desert Extreme	NA	NA	Sunny	Cool
22-Oct	2	DESERT EXTREME	1010	1240	150	COLLECTING DATA	Collecting Data, North - South, West - East	GPS	Linear	Sunny	Cool
22-Oct	2	DESERT EXTREME	1240	1330	50	DEMOBILIZATION	End of test< packing up	NA	NA	Sunny	Warm

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

APPENDIX E. REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

4. Yuma Proving Ground Soil Survey Report, May 2003.

5. Practical Nonparametric Statistics, W.J. Conover, John Wiley & Sons, 1980, pages 144 through 151.

APPENDIX F. ABBREVIATIONS

ADST	=	Aberdeen Data Services Team
APG	=	Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
ATSS	=	Aberdeen Test Support Services
E	=	efficiency
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
EQT	=	Army Environmental Quality Technology Program
GPS	=	Global Positioning System
HDSD	=	Homeland Defense and Sustainment Division
HEAT	=	high-explosive antitank
JPG	=	Jefferson Proving Ground
M	=	standard deviation
NS	=	nonstandard
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver-operating characteristic
SERDP	=	Strategic Environmental Research and Development Program
SL	=	Survivability and Lethality
USAEC	=	U.S. Army Environmental Command
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground

APPENDIX G. DISTRIBUTION LIST

DTC Project No.8-CO-160-UXO-021

Addressee	No. of Copies
Commander U.S. Army Aberdeen Test Center ATTN: TEDT-AT-SLE (Mr. J. Stephen McClung) 400 Colleran Road Aberdeen Proving Ground, MD 21005-5059	1
Program Manager SERDP/ESTCP Munitions Management ATTN: Mr. Herb Nelson 901 North Stuart Street, Suite 303 Arlington, VA 22203	1
Sky Research Inc. ATTN: Dr. Nicolas Lhomme 44 Dead Indian Memorial Road Ashland, Oregon 97520	1
Defense Technical Information Center 8725 John J. Kingman Road, STE 0944 Fort Belvoir, VA 22060-6218	PDF

Secondary distribution is controlled by Program Manager, SERDP/ESTCP, Munitions Management, ATTN: Mr. Herb Nelson.